

On the maximum entropy and the problem of moments. An application to stellar kinematics

Rafael Cubarsi

Dept. Matemàtica Aplicada IV
Universitat Politècnica de Catalunya
08034 Barcelona, Catalonia, Spain
E-mail: rcubarsi@ma4.upc.edu

December 2008

Abstract

The maximum entropy approach is used to solve the classical moment problem of stellar kinematics. If an extended set of moments is available, the current method provides a linear estimation algorithm, which is given by a Gramian system of equations, that leads to a fast and suitable estimation of the velocity distribution. In particular, it can be used as an alternative approach for modelling multimodal distributions that can not be described through gaussian mixtures. Two samples drawn from the Hipparcos Catalog are studied. The first one, which is composed of a non-gaussian mixture of early-type and young disk stars, is used to test the ability of the procedure in reconstructing a truncated distribution associated with a bounded stellar sample. By fitting up to tenth moments, the maximum entropy approach gives a realistic portrait of its actual asymmetries. The second one is a nearly complete and kinematically representative sample of the solar neighbourhood, containing thin and thick disk populations. The local velocity distribution can be described by fitting moments up to sixth-order as a product of two exponential functions in the form $f = \varphi_1(Q) \varphi_2(h)$, where Q is a quadratic positive definite form, which gives the background ellipsoidal shape of the distribution, and the function $\varphi_2(h)$, which can be written in terms of the angular momentum integral h , is a perturbation factor, which is quadratic in the V velocity alone, and gives account of the skewness and of the slight shift of the velocity ellipsoids in terms of the rotation velocity.

KEY WORDS: stars: kinematics – galaxies: kinematics and dynamics – galaxies: statistics – methods: statistical.

1991 MATHEMATICS SUBJECT CLASSIFICATION: 60, 62, 85.

1 Introduction

The asymmetry of the local velocity distribution was first studied in 1905 by Kapteyn in his theory of two star streams, and it was further developed by Kapteyn (1922), Strömberg (1925) and Charlier (1926), which took into account up to fourth moments of the velocity distribution. However, those moments were not determined with a sufficient degree of accuracy up to Erickson (1975). During the last decade, higher-order velocity moments with better precision have been obtained from the Hipparcos Catalog (ESA 1997), allowing to give account for velocity discontinuities and kinematic populations in the solar neighbourhood (Cubarsi & Alcobé 2004, Alcobé & Cubarsi, 2005). Several approaches have been tried to describe the asymmetry of the velocity distribution. For example, it may be expressed as a mixture of stellar populations with trivariate gaussian partial distributions (Soubiran & Girard 2005, Vallerani et al. 2006). Alternatively, a two- or three-integral model based in Fricke (1952) components (Evans et al. 1997, Famaey et al. 2002, Jiang & Ossipkov 2007), or even a combination of a gaussian part of the density function with a perturbation factor expressed in a polynomial form in terms of the integrals of motion may be used (van der Marel & Franx 1993, Gerhard 1993, Kormendy et al. 1998). Although to this purpose the velocity distribution is sometimes numerically estimated (Dehnen 1998, Skuljan et al. 1999), it is also frequent the use of analytical models (Famaey et al. 2005, Veltz et al. 2008). However, in the later case, according to todays observational data, and depending on the statistical model, some intricate trivariate distribution functions may be obtained. In the present work, an alternative analytical model based on the maximum entropy approach is proposed to explain the eventual asymmetric features of the velocity distribution, which are shown through the population moments. Even though such an approach has been widely used to solve many technical and scientific problems, to my knowledge, there is not any application to stellar kinematics. Due to its simplicity it could be worthwhile to use this method to construct any ad hoc velocity distribution function.

In the following sections, the necessary complexity of the velocity distribution to satisfy a set of moment constraints will be studied. The purpose of the current approach is to simplify as much as possible both, analytical dependence and parameter estimation, of the distribution function. This will be done under the following circumstances.

1. A density function maximising Shannon's information entropy will be chosen. The maximum entropy approach to the solution of inverse problems was introduced long time ago by Jaynes (1957), so that it provides a uniquely solution which is maximally noncommittal with regard to missing information, and that it agrees with what is known, but express maximum uncertainty with respect to all other matters. It is a flexible and powerful tool for density approximation, which nests a whole family of generalised exponential distributions, including the exponential, normal, lognormal, gamma, beta as special cases. Other properties of maximum entropy distributions will be given in section §2.

An interesting application of the maximum entropy approach is the problem of moments (Mead & Papanicolaou 1984), which we describe below, while introducing the notation according to the astronomical formulation,

For fixed values of time t and position \mathbf{r} , the macroscopic properties of a stellar system can be described from the moments of the distribution, which provide indirect information of the phase space density function $f(t, \mathbf{r}, \mathbf{V})$, that will be assumed as normalised in regard to the velocities. Then, the symmetric tensor of the n^{th} -order non-centred trivariate moments

is obtained from the following expected value

$$\mathbf{m}_n(t, \mathbf{r}) = \langle (\mathbf{V})^n \rangle \equiv \int_{\Gamma_V} (\mathbf{V})^n f(t, \mathbf{r}, \mathbf{V}) d\mathbf{V}, \quad n \geq 0 \quad (1)$$

where $(\cdot)^n$ stands for the n^{th} -tensor power, and Γ_V is the velocity space domain. The tensor \mathbf{m}_n has $\binom{n+2}{2}$ different elements according to the expression

$$m_{i_1 i_2 \dots i_n} = \langle V_{i_1} V_{i_2} \dots V_{i_n} \rangle \quad (2)$$

so that the indices belong to the set $\{1, 2, 3\}$, depending on the velocity component. Sometimes, instead of the component notation, namely in latin indices, the notation making explicit the velocity powers is used, namely in greek indices, according to

$$m_{\alpha\beta\gamma} = \langle V_1^\alpha V_2^\beta V_3^\gamma \rangle \quad (3)$$

Obviously, $m_0 = 1$, and $\mathbf{m}_1 = \mathbf{v}(t, \mathbf{r})$ is the mean velocity, or velocity of the centroid. In a similar way, the symmetric tensor of the n^{th} -order centred moments can be obtained by working from the peculiar velocity

$$\mathbf{u} = \mathbf{V} - \mathbf{v}(t, \mathbf{r}) \quad (4)$$

as $\mu_n = \langle (\mathbf{u})^n \rangle$, with elements $\mu_{i_1 i_2 \dots i_n}$. In this case, $\mu_0 = 1$ and $\mu_1 = \mathbf{0}$.

Hereafter, when studying from an statistical viewpoint the velocity dependence of the distribution function, the variables of time and position will be omitted, although they might be used in the framework of a dynamical model for the whole phase space distribution function.

Provided an ordination of a set of moments, for example according to indices 0, 1, 2, 3, 11, 12, 13, 22, and so on, if the first m moments are known, it is possible to find an infinite variety of functions whose first m moments coincide with the above set. Various approximation procedures exist in order to find a sequence of functions f_m , which fulfils the foregoing moment constraints and converges to the true distribution as m approaches infinity. Fortunately, between those sequences of functions there exists a uniquely maximum entropy sequence which maximise the entropy

$$W(f_m) = - \int_{\Gamma_V} f_m(\mathbf{V}) \ln f_m(\mathbf{V}) d\mathbf{V} \quad (5)$$

Then, the maxima $f = f_m$ is usually called the least biassed sequence of approximations, and, by using Lagrangian multipliers, it can be shown (e.g. Kagan et al. 1973) that it has the form

$$f(\mathbf{V}) = e^{\mathcal{P}(\mathbf{V})} \quad (6)$$

where $\mathcal{P}(\mathbf{V})$ is a power series of the velocity components containing m terms, as many terms as the number of moment constraints, so that each coefficient is related to a single moment constraint. Then, the solution of the maximum entropy problem consists in to solve a set of m non-linear equations in the form

$$\mathbf{m}_k = \int_{\Gamma_V} (\mathbf{V})^k e^{\mathcal{P}(\mathbf{V})} d\mathbf{V} \quad (7)$$

However, a couple of remarks must be done. On one hand, these techniques are typically not easy to generalise to the non-unidimensional problem. On the other hand, even for the unidimensional problem, an analytical solution does not generally exist for higher than second moments. Generally, the numerical techniques for solving the coefficients of the polynomial \mathcal{P} are based on nonlinear optimisation, Legendre transformation, etc. (e.g. de Bruin et al. 1999, Kouskoulas et al. 2004) and, in any case, they are not computationally trivial. However, if an extended set of moments is known –in the case of trivariate distributions, for a polynomial of degree n in the three variables it is necessary to compute moments up to order $2(n-1)$ – then the parameter estimation can be done linearly, as it is shown in section §4.

2. The current purpose is to infer the trivariate velocity distribution from a finite set of moment constraints. In order to simplify the estimation of the above polynomial coefficients, an alternative method has been developed, which is based in an unique assumption: The velocity distribution satisfies the boundary conditions associated with the moment equations.

Indeed, if the phase space distribution function f satisfies the collisionless Boltzmann equation, $\frac{Df}{Dt} = 0$, then by multiplying it by the n^{th} -tensor power of the star velocity, and by integrating over the whole velocity space, the family of stellar hydrodynamic equations can be obtained

$$\int_{\Gamma_V} (\mathbf{V})^n \frac{Df}{Dt} d^3\mathbf{V} = (\mathbf{0})^n, \quad n \geq 0 \quad (8)$$

In Cubarsi (2007) above equations were derived in terms of the central velocity moments, in a completely analytical way, for any order n , and without any additional hypotheses. Then, if above integrals exist, in the integration process, since there are not stars with velocity beyond Γ_V , the following boundary conditions were, as usual, assumed,

$$(\mathbf{V})^n f(t, \mathbf{r}, \mathbf{V}) |_{\mathbf{V} \in \partial\Gamma_V} = (\mathbf{0})^n, \quad n \geq 0 \quad (9)$$

One of the integral properties that was derived in Cubarsi (2007) to another purpose, will now allow, in section §3, to establish a Gramian system of equations associated with our estimation problem.

Notice that, from a purely statistical inference viewpoint, the requirement to estimate the distribution parameters is not that the phase density function is solution of the collisionless Boltzmann equation, but it is sufficient that it satisfies above boundary conditions. Let us also note that the entropy functional $W(f)$, as defined in Eq. 5, is far from containing all the information about the Boltzmann equation, since $W(f)$ depends only on the velocity space, similarly to the collision operator of the Boltzmann equation. In the following section we shall discuss how such a maximum entropy density function may be a solution of the collisionless Boltzmann equation.

Let us review two typical cases of maximum entropy distribution function, which are solution of the whole set of moment equations. The simplest case is an isothermal velocity distribution of Maxwell type in the peculiar velocities, which, according to the Maxwell-Boltzmann law, represents a system with the more basic thermal equilibrium.

$$\psi(t, \mathbf{r}, \mathbf{u}) = e^{-\frac{1}{2}\mu^{-1}|\mathbf{u}|^2} \quad (10)$$

where $\mu(t, \mathbf{r}) > 0$ is a continuous and differentiable function in both arguments, giving account of the variance of the distribution. However, since this distribution is totally isotropic,

it has equal diagonal second central moments, vanishing off-diagonal second moments, and null odd-order moments as well.

Another well known example is the Schwarzschild distribution, that is, an exponential density function depending on the peculiar velocities in a quadratic way (Chandrasekhar 1942),

$$\psi(t, \mathbf{r}, \mathbf{u}) = e^{-\frac{1}{2}(Q+\sigma)}, \quad Q = \mathbf{u}^T \cdot \mathbf{A}_2 \cdot \mathbf{u} \quad (11)$$

where Q is a quadratic, positive definite form, with $\mathbf{A}_2(t, \mathbf{r})$ a second-rank symmetric tensor and $\sigma(t, \mathbf{r})$ a scalar function, which are continuous and differentiable in both arguments. Hence, the distribution is of gaussian type in the peculiar velocities, although it is multiplied by an arbitrary function of time and position. In such a way, the quadratic form Q can give account of three isolating integrals of star motions, so that, in general, the distribution may have some different diagonal second central moments and non-vanishing off-diagonal moments, although the odd-order moments still vanishes.

Above examples, which are integrable functions in an infinite velocity domain, satisfy the boundary conditions, Eq. 9, and can be generalised according to an exponential function, Eq. 6, with as many polynomial terms as available moments, under the necessary conditions over the polynomial coefficients in order to obtain an integrable distribution function. For higher-degree polynomials, the distribution function is integrable if the polynomial is upper bounded.

For truncated distributions which are associated with velocity bounded stellar samples, e.g. $|\mathbf{V} - \mathbf{V}_0| \leq \text{const}$, we should consider a finite velocity domain. Then the boundary conditions are also valid if, for any of the velocity components V_i , the values of $V_i^n e^{\mathcal{P}(\mathbf{V})}$ are similar in both extremes of the domain Γ_{V_i} of this velocity component. This is fulfilled, for example, if the distribution is symmetric with respect to one of the variables, like in the case of disk samples. Though, even in a case where the boundary conditions were not exactly fulfilled, they would be still a good approximation if the truncated distribution nearly vanishes as approaching the contour of the velocity domain. Out of this boundary, the density function must be assumed null. Thus, for a domain either bounded or unbounded, we shall assume that the velocity distribution is continuous, differentiable and positive in the interior of the velocity domain Γ_V , and that the boundary conditions are fulfilled in its contour $\partial\Gamma_V$.

In section §5 two case examples illustrate the maximum entropy approach for two local stellar samples drawn from the Hipparcos Catalogue (ESA 1997). For testing purposes a truncated distribution is studied, which is associated with a velocity bounded sample (Sample I) containing, basically, early-type and young disk stars. Its velocity distribution needs a six degree polynomial to describe its strong asymmetries, where moments up to order ten are fitted. The other application shows the basic features of the local distribution working from a kinematically representative sample (Sample II) containing thin and thick disk stars. Its velocity distribution is accurately described from a four degree polynomial, by fitting the moments up to sixth-order. In both cases the entropy approach gives a consistent portrait of the distribution, as it is discussed in the last section.

2 Maximum entropy

By quoting Jaworsky (1987), there are two typical viewpoints for interpreting the entropy as uncertainty. In mathematical statistics and information theory the entropy functional is maximised

attending to some constraints that express any available information of a complex physical system, which depend on the actual experimental situation. In statistical mechanics the entropy is used to study the thermodynamic equilibrium or non-equilibrium of a physical system, generally a uniform gas, in terms of the mean values of some physical quantities, which describe the macroscopic state of a physical system as a whole, like energy or number of particles. Thus, statistical mechanics based on this principle can be interpreted as a special type of statistical inference. The use of higher-order statistical moments in addition to the mean values represents a generalisation of the thermodynamic concept of entropy, which is used to approximate the exact probability distributions for a few specified random variables when a finite number of their moments is known, although, from a purely thermodynamic point of view, the extra information provided by the higher-order moments has been proved non-essential.

In galactic dynamics, the phase density function, due to its physical significance, should satisfy the collisionless Boltzmann equation. The clue to answer the question of how a maximum entropy function, as Eq. 5, should satisfy such an equation may arise by reviewing both examples of the above section. For elastic collisions involving short range forces, and in absence of boundaries, mass, momentum, and energy are conserved under binary encounters (e.g. Cercignani 1988). They are usually referred as collisional invariants. As it is well known, there is only one distribution function, the Maxwellian distribution, owing all of the following properties: it depends on a linear combination of the collisional invariants, the collision term of the Boltzmann equation is exactly zero, and it is a maximum entropy function. This solution represents a local equilibrium state, in the sense that other solutions to the Boltzmann equation will become closer to it as the time goes by. However, depending on the boundary conditions, and on the dissipative effects, maximum entropy solutions can be non-Maxwellian, like in Lynden-Bell (1967) for a rotating stellar system, which is a particular case of the Schwarzschild distribution. Hence, in general, there exist more complex distributions than the Maxwellian which are solution of the collisionless Boltzmann equation, and they are maximum entropy distributions, such as a trivariate Schwarzschild distribution, which are commonly accepted as a functional approach to describe some stellar systems. Indeed, they may be obtained by assuming that the Liouville's theorem is satisfied, so that the essential information about the density function is provided by the isolating integrals of the motion of the stars. Thus, if we assume that the polynomial form \mathcal{P} of Eq. 5 depends on the integrals of motion, and is itself an integral of motion, the Liouville's theorem is equivalent to the collisionless Boltzmann approximation. Then, the collisionless Boltzmann equation obviously takes the form

$$\frac{df(\mathcal{P})}{d\mathcal{P}} \frac{D\mathcal{P}}{Dt} = 0 \quad (12)$$

so that the factor $\frac{df}{d\mathcal{P}}$ accounts for the maximum entropy condition, and the factor $\frac{D\mathcal{P}}{Dt}$ is, in fact, the collisionless Boltzmann condition. Thus, both conditions are independent and compatible. If the maximum entropy criterion is fulfilled, then $f(\mathcal{P}) = e^{\mathcal{P}}$ takes the smoothest possible form, while the dependence of \mathcal{P} in terms of the powers of the velocity, as well as in terms of time and position through its polynomial coefficients, is, under this approach, independent from the maximum entropy condition. Therefore, we can also affirm that the maximum entropy procedure is non-essential to the solution of the collisionless Boltzmann equation.

Let us then write the maximum entropy density function as

$$f = \psi_n \equiv e^{\mathcal{P}_n}, \quad \mathcal{P}_n = \sum_{k=0}^n \sum_{\alpha+\beta+\gamma=k} \lambda_{\alpha\beta\gamma}^{(k)}(t, \mathbf{r}) V_1^\alpha V_2^\beta V_3^\gamma \quad (13)$$

where the subindex n does not represent the number of polynomial terms, but the maximum polynomial power.

If the velocity domain Γ_V is all the space \mathbb{R}^3 , the polynomial \mathcal{P}_n must be upper bounded in order to satisfy the integrability conditions. Hence the power series of the velocities reaches a natural value n , which must be *even*, and for the highest degree, $k = n$, $\sum_{\alpha+\beta+\gamma=n} \lambda_{\alpha\beta\gamma}^{(n)}(t, \mathbf{r}) V_1^\alpha V_2^\beta V_3^\gamma$ must be a negative definite n -adic form.

We also need to write \mathcal{P}_n of Eq. 13 by using a slightly different notation, with latin indices instead of greek indices, so that each term accounts for products of the same degree in the velocities. Einstein's summation criterion for repeated indices will be hereafter used. Thus,

$$\mathcal{P}_n = \lambda_0 + \lambda_i V_i + \lambda_{ij} V_i V_j + \cdots + \lambda_{i_1 \dots i_n} V_{i_1} \cdots V_{i_n} \quad (14)$$

In the term corresponding to the k th-power of the velocities, since the coefficients $\lambda_{i_1 \dots i_k}$ are symmetric, we have $\binom{k+2}{2} = \binom{k+2}{k}$ different coefficients. Hence, up to the n th-power there are $\sum_{k=0}^n \binom{k+2}{k} = \binom{n+3}{n}$ different coefficients. In addition, we shall use the following relationship

$$\lambda_{\alpha\beta\gamma}^{(k)} = \frac{k!}{\alpha!\beta!\gamma!} \lambda_{\underbrace{1\dots 1}_\alpha \underbrace{2\dots 2}_\beta \underbrace{3\dots 3}_\gamma}, \quad k = \alpha + \beta + \gamma \quad (15)$$

which establishes the correspondence between the greek and latin indices notations for the coefficients of \mathcal{P}_n .

Some aspects of the maximum entropy distribution function may be still pointed out:

Eq. 13, in addition to include as particular cases Eq. 10 and Eq. 11, it also contains, in general, any desired type of two- or three-integral functions (e.g. Hénon 1973, Dejonghe 1983, White 1985). It represents a quite general functional approach, in a similar way as Fricke (1952), with the difference that, while in the Fricke-based models the distribution function is either linear combination or product of powers of the integrals of motion, in Eq. 13 the linear combination of powers of integrals of motion appears as argument of the exponential function.

When $n \rightarrow \infty$, Eq. 13 converges to the true distribution. Then, if the velocity distribution is expressed as a power series of the velocities, we have

$$\begin{aligned} f(t, \mathbf{r}, \mathbf{V}) &= c_0 \left(1 + \sum_{k=1}^{\infty} \sum_{\alpha+\beta+\gamma=k} c_{\alpha\beta\gamma}^{(k)}(t, \mathbf{r}) V_1^\alpha V_2^\beta V_3^\gamma \right) = \\ &= \exp \left(\sum_{k=0}^{\infty} \sum_{\alpha+\beta+\gamma=k} \lambda_{\alpha\beta\gamma}^{(k)}(t, \mathbf{r}) V_1^\alpha V_2^\beta V_3^\gamma \right) \end{aligned} \quad (16)$$

which is the relationship between generalised moments and cumulants (Stuart & Ord 1987), where the coefficient $c_0 = \exp(\lambda_0^{(0)})$ provides the normalisation of the distribution. Let us remember that, when a similar relation holds for the characteristic function $\Phi(\xi)$, which is the Fourier transform of $f(\mathbf{V})$, then the coefficients $c_{\alpha\beta\gamma}^{(k)}$ become proportional to the population moments $m_{\alpha\beta\gamma}$, and $\lambda_{\alpha\beta\gamma}^{(k)}$ become proportional to the cumulants of the distribution $\kappa_{\alpha\beta\gamma}$, by a factor $\frac{1}{\alpha!\beta!\gamma!}$.

A maximum entropy distribution function can exhibit several modes. In the trivariate case, if Eq. 13 has a polynomial of even degree n , the distribution can exhibit $(n/2)^3$ modes, since an univariate exponential with a polynomial of degree n may have up to $n/2$ modes. In general, it is necessary to estimate less number of parameters for Eq. 13 than for a mixture of trivariate gaussian distributions giving account for the same number of modes.

In addition, under maximum entropy distributions, the sample moments are maximum likelihood estimators of the population moments.

3 Boundary conditions

We study a quite general case of fitting a defined set of velocity moments, up to order $2(n-1)$, with a maximum entropy velocity distribution containing a polynomial of degree n , which allows a simple and linear estimation of the polynomial coefficients. By using latin indices notation for \mathcal{P}_n , according to Eq. 14, we assume that all the moments in the form of Eq. 43 exist, what is equivalent to consider the distribution function to be a square-integrable function in the velocity domain Γ_V . The scalar λ_0 is the normalisation factor, and, in general, all of the above coefficients are symmetric elements of the k -rank tensors λ_k , $k = 0, \dots, n$.

The coefficients, other than λ_0 , can be obtained by using the following property, which, in particular, is fulfilled by any solution of the moment equations,

$$\int_{\Gamma_V} \nabla_{\mathbf{V}} [(\mathbf{V})^m e^{\mathcal{P}_n}] d^3\mathbf{V} = (\mathbf{0})^{n+1} \quad (17)$$

Above integral is a $(n+1)$ -rank tensor, which is symmetric with respect to the indices of the tensor power $(\mathbf{V})^m$. Eq. 17 is a direct consequence of Eq. 9. Thus, when integrating Eq. 17 by components, and the conditions of Eq. 9 are applied over the domain of the variable $V_{i_{m+1}}$, we get

$$\int_{\Gamma_{V_{i_{m+1}}}} \frac{\partial(V_{i_1} \dots V_{i_m} e^{\mathcal{P}_n})}{\partial V_{i_{m+1}}} dV_{i_{m+1}} = V_{i_1} \dots V_{i_m} e^{\mathcal{P}_n}|_{\Gamma_{V_{i_{m+1}}}} = 0 \quad (18)$$

In particular, for $m = 0$, since

$$\frac{\partial e^{\mathcal{P}_n}}{\partial V_k} = e^{\mathcal{P}_n} \frac{\partial \mathcal{P}_n}{\partial V_k} \quad (19)$$

we have

$$\int_{\Gamma_V} \frac{\partial \mathcal{P}_n}{\partial V_k} e^{\mathcal{P}_n} d^3\mathbf{V} = 0 \quad (20)$$

Similarly, for $m = 1$,

$$\int_{\Gamma_V} \frac{\partial}{\partial V_k} (V_i e^{\mathcal{P}_n}) d^3\mathbf{V} = \int_{\Gamma_V} \delta_{ik} e^{\mathcal{P}_n} d^3\mathbf{V} + \int_{\Gamma_V} V_i \frac{\partial \mathcal{P}_n}{\partial V_k} e^{\mathcal{P}_n} d^3\mathbf{V} = 0 \quad (21)$$

where δ_{ik} is the Kronecker delta.

And, in general, for $m \geq 2$, we get

$$\begin{aligned} & \frac{\partial(V_{i_1} \dots V_{i_m} e^{\mathcal{P}_n})}{\partial V_{i_{m+1}}} = \\ & = \left(\delta_{i_1 i_{m+1}} V_{i_2} \dots V_{i_m} + \dots + \delta_{i_j i_{m+1}} V_{i_1} \dots \widehat{V_{i_j}} \dots V_{i_m} + \dots + \delta_{i_m i_{m+1}} V_{i_1} \dots V_{i_{m-1}} \right) e^{\mathcal{P}_n} + \\ & + V_{i_1} \dots V_{i_m} \frac{\partial e^{\mathcal{P}_n}}{\partial V_{i_{m+1}}} \end{aligned} \quad (22)$$

where the hat remarks the omitted factors. Once more, bearing in mind Eq. 19, the identity Eq. 17 yields

$$\begin{aligned} & \int_{\Gamma_V} \left(\delta_{i_1 i_{m+1}} V_{i_2} \dots V_{i_m} + \dots + \delta_{i_j i_{m+1}} V_{i_1} \dots \widehat{V_{i_j}} \dots V_{i_m} + \dots + \delta_{i_m i_{m+1}} V_{i_1} \dots V_{i_{m-1}} \right) e^{\mathcal{P}_n} d^3\mathbf{V} + \\ & + \int_{\Gamma_V} V_{i_1} \dots V_{i_m} \frac{\partial \mathcal{P}_n}{\partial V_{i_{m+1}}} e^{\mathcal{P}_n} d^3\mathbf{V} = 0 \end{aligned} \quad (23)$$

Notice that, since the first integral is symmetric with respect to permutation of indices, and, in general, it is not null, then the second integral

$$\mathbf{q}_{m+1} = \int_{\Gamma_V} (\mathbf{V})^m \otimes (\nabla_{\mathbf{V}} \mathcal{P}_n) e^{\mathcal{P}_n} d^3 \mathbf{V} \quad (24)$$

must be symmetric too.

Indeed, Eq. 23 and Eq. 24 are equivalent to those obtained in Cubarsi (2007) as Eq. 22 and Eq. 29, which were derived to express the conservation of pressures.

Now, above identities will provide a linear method in order to fit any desired set of moments. In contrast with the usual maximum entropy methods for the moments problem, which are non linear and not enough well conditioned, the present method allows to determine with accuracy all of the coefficients. On the other hand, the family of distributions satisfying the boundary conditions, Eq. 9, is a quite general one.

First we write $\nabla_{\mathbf{V}} \mathcal{P}_n$ from Eq. 14,

$$\frac{\partial \mathcal{P}_n}{\partial V_k} = \lambda_k + 2\lambda_{j_1 k} V_{j_1} + 3\lambda_{j_1 j_2 k} V_{j_1} V_{j_2} + \cdots + n\lambda_{j_1 j_2 \dots j_{n-1} k} V_{j_1} \cdots V_{j_{n-1}} \quad (25)$$

In order to obtain all of the elements of tensors λ_k , $k = 1, \dots, n$, we compute the integrals of Eq. 17 for m from 0 to $n - 1$.

For $m = 0$, by taking into account Eq. 20 and Eq. 25, and by using the moments definition, Eq. 1, since $m_0 = 1$, we have,

$$\lambda_k + 2\lambda_{j_1 k} m_{j_1} + 3\lambda_{j_1 j_2 k} m_{j_1 j_2} + \cdots + n\lambda_{j_1 j_2 \dots j_{n-1} k} m_{j_1 \dots j_{n-1}} = 0 \quad (26)$$

which stands for a set of 3 scalar equations, $k = 1, 2, 3$.

For $m = 1$, also by taking into account Eq. 21 and Eq. 25, we get

$$\delta_{ik} + \lambda_k m_i + 2\lambda_{j_1 k} m_{j_1 i} + 3\lambda_{j_1 j_2 k} m_{j_1 j_2 i} + \cdots + n\lambda_{j_1 j_2 \dots j_{n-1} k} m_{j_1 j_2 \dots j_{n-1} i} = 0 \quad (27)$$

Hence, this set of relations, for $i, k = 1, 2, 3$, due to the symmetry of Eq. 24, provides 6 independent scalar equations.

And, in general, for $m = n - 1$, from Eq. 23 we similarly get

$$\begin{aligned} & \delta_{i_1 i_n} m_{i_2 \dots i_{n-1}} + \cdots + \delta_{i_j i_n} m_{i_1 \dots \widehat{i_j} \dots i_{n-1}} + \cdots + \delta_{i_{n-1} i_n} m_{i_1 \dots i_{n-2}} + \\ & + \lambda_{i_n} m_{i_1 \dots i_{n-1}} + 2\lambda_{j_1 i_n} m_{j_1 i_1 \dots i_{n-1}} + 3\lambda_{j_1 j_2 i_n} m_{j_1 j_2 i_1 \dots i_{n-1}} + \cdots + n\lambda_{j_1 j_2 \dots j_{n-1} i_n} m_{j_1 \dots j_{n-1} i_1 \dots i_{n-1}} = 0 \end{aligned} \quad (28)$$

which consists, owing to the symmetry of Eq. 24, in a set of $\binom{n+2}{2}$ independent scalar equations, for $i_1, \dots, i_n = 1, 2, 3$.

Therefore, we have as many linear equations as unknowns composing the elements of symmetric tensors λ_k , $k = 1, \dots, n$, whose elements are the coefficients of \mathcal{P}_n . Such a non-homogeneous system can be associated with a Gramian matrix, as shown in the next section.

Finally, the scalar λ_0 , lasting to be evaluated, may simply obtained as the normalisation factor in order to satisfy

$$\int_{\Gamma_V} e^{\mathcal{P}_n} d^3 \mathbf{V} = 1 \implies e^{-\lambda_0} = \int_{\Gamma_V} e^{\lambda_i V_i + \lambda_{ij} V_i V_j + \cdots + \lambda_{i_1 \dots i_n} V_{i_1} \cdots V_{i_n}} d^3 \mathbf{V} \quad (29)$$

	0	1	2	3	11	12	13	22	23	33	111	112	...	333	...
0	1	m_1	m_2	m_3	m_{11}	m_{12}	m_{13}	m_{22}	m_{23}	m_{33}	m_{111}	m_{112}	...	m_{333}	...
1		m_{11}	m_{12}	m_{13}	m_{111}	m_{112}	m_{113}	m_{122}	m_{123}	m_{133}	m_{1111}	m_{1112}	...	m_{1333}	...
2			m_{22}	m_{23}	m_{112}	m_{122}	m_{123}	m_{222}	m_{223}	m_{233}	m_{1112}	m_{1122}	...	m_{2333}	...
3				m_{33}	m_{113}	m_{123}	m_{133}	m_{223}	m_{233}	m_{333}	m_{1113}	m_{1123}	...	m_{3333}	...
11					m_{1111}	m_{1112}	m_{1113}	m_{1122}	m_{1123}	m_{1133}	m_{11111}	m_{11112}	...	m_{111333}	...
12						m_{1122}	m_{1123}	m_{1222}	m_{1223}	m_{1233}	m_{11112}	m_{11122}	...	m_{12333}	...
13							m_{1133}	m_{1223}	m_{1233}	m_{1333}	m_{11113}	m_{11123}	...	m_{13333}	...
22								m_{2222}	m_{2223}	m_{2233}	m_{11122}	m_{11222}	...	m_{22333}	...
23									m_{2233}	m_{2333}	m_{11123}	m_{11223}	...	m_{23333}	...
33										m_{3333}	m_{11133}	m_{11233}	...	m_{33333}	...
111											m_{111111}	m_{111112}	...	$m_{1111333}$...
112												m_{111122}	...	$m_{1112333}$...
...												
333														m_{333333}	...
...															...

(31)

Table 1: Matrix \mathbf{G}_2 , which is a symmetric matrix of inner products of the velocity components $\langle V_0 V_i V_j \dots, V_0 V_p V_q \dots \rangle$, according to latin indices, with $V_0 \equiv 1$ and the other indices sorted as $1 \leq i \leq j \leq \dots \leq 3$ and $1 \leq p \leq q \leq \dots \leq 3$. The first row and first column are referring to the velocity indices. Since the matrix is symmetric, only the diagonal and upper triangular part are written.

4 Gramian system

The three scalar equations involved in Eq. 26, corresponding to $m = 0$, for $k = 1, 2, 3$, are homogeneous in the elements of tensors λ_k . In Eq. 27, for $m = 1$, we group the terms containing the elements of λ_k , by writing the other ones in the right hand side. And similarly, for the general equation with $m = n - 1$, Eq. 28. Thus we obtain the following linear system of equations for the elements of tensors λ_k ,

$$\begin{aligned}
\lambda_k + 2\lambda_{j_1 k} m_{j_1} + 3\lambda_{j_1 j_2 k} m_{j_1 j_2} + \dots + n\lambda_{j_1 j_2 \dots j_{n-1} k} m_{j_1 \dots j_{n-1}} &= 0 \\
\lambda_k m_i + 2\lambda_{j_1 k} m_{j_1 i} + 3\lambda_{j_1 j_2 k} m_{j_1 j_2 i} + \dots + n\lambda_{j_1 j_2 \dots j_{n-1} k} m_{j_1 j_2 \dots j_{n-1} i} &= -\delta_{ik} \\
\dots &
\end{aligned}
\tag{30}$$

$$\begin{aligned}
&\lambda_{i_n} m_{i_1 \dots i_{n-1}} + 2\lambda_{j_1 i_n} m_{j_1 i_1 \dots i_{n-1}} + 3\lambda_{j_1 j_2 i_n} m_{j_1 j_2 i_1 \dots i_{n-1}} + \dots + n\lambda_{j_1 j_2 \dots j_{n-1} i_n} m_{j_1 \dots j_{n-1} i_1 \dots i_{n-1}} = \\
&= -\left(\delta_{i_1 i_n} m_{i_2 \dots i_{n-1}} + \dots + \delta_{i_j i_n} m_{i_1 \dots \widehat{i_j} \dots i_{n-1}} + \dots + \delta_{i_{n-1} i_n} m_{i_1 \dots i_{n-2}} \right)
\end{aligned}$$

Such a system of equations can be grouped according to three different vectors, in regard to their right hand side, for $k = 1, 2, 3$ in the first two equations, and for $i_n = 1, 2, 3$ in the general expression. A similar procedure can be applied to the λ_k coefficients.

The system matrix, namely \mathbf{G}_2 , as shown in Table 1, can be interpreted as a symmetric matrix of inner products of the velocity components $\langle V_0 V_i V_j \dots, V_0 V_p V_q \dots \rangle$ with respect to the weight function $e^{\mathcal{P}_n}$, according to latin indices notation, with $V_0 \equiv 1$ and the other indices sorted as $1 \leq i \leq j \leq \dots \leq 3$ and $1 \leq p \leq q \leq \dots \leq 3$.

Therefore, \mathbf{G}_2 is a Gram matrix, symmetric, positive definite and, among other well known properties, it is invertible. Hence, the system has a unique solution.

	A	B	C	a	b	c
$m = 0$	0	0	0	λ_1	λ_2	λ_3
$m = 1$	-1	0	0	$2\lambda_{11}$	$2\lambda_{21}$	$2\lambda_{31}$
	0	-1	0	$2\lambda_{12}$	$2\lambda_{22}$	$2\lambda_{32}$
	0	0	-1	$2\lambda_{13}$	$2\lambda_{23}$	$2\lambda_{33}$
$m = 2$	$-2 m_1$	0	0	$3 \lambda_{111}$	$3 \lambda_{211}$	$3 \lambda_{311}$
	$-m_2$	$-m_1$	0	$2 \cdot 3 \lambda_{112}$	$2 \cdot 3 \lambda_{212}$	$2 \cdot 3 \lambda_{312}$
	$-m_3$	0	$-m_1$	$2 \cdot 3 \lambda_{113}$	$2 \cdot 3 \lambda_{213}$	$2 \cdot 3 \lambda_{313}$
	0	$-2 m_2$	0	$3 \lambda_{122}$	$3 \lambda_{222}$	$3 \lambda_{322}$
	0	$-m_3$	$-m_2$	$2 \cdot 3 \lambda_{123}$	$2 \cdot 3 \lambda_{223}$	$2 \cdot 3 \lambda_{323}$
	0	0	$-2 m_3$	$3 \lambda_{133}$	$3 \lambda_{233}$	$3 \lambda_{333}$
$m = 3$	$-3 m_{11}$	0	0	$4 \lambda_{1111}$	$4 \lambda_{2111}$	$4 \lambda_{3111}$
	$-2 m_{12}$	$-m_{11}$	0	$3 \cdot 4 \lambda_{1112}$	$3 \cdot 4 \lambda_{2112}$	$3 \cdot 4 \lambda_{3112}$
	$-2 m_{13}$	0	$-m_{11}$	$3 \cdot 4 \lambda_{1113}$	$3 \cdot 4 \lambda_{2113}$	$3 \cdot 4 \lambda_{3113}$
	$-m_{22}$	$-2 m_{12}$	0	$3 \cdot 4 \lambda_{1122}$	$3 \cdot 4 \lambda_{2122}$	$3 \cdot 4 \lambda_{3122}$
	$-m_{23}$	$-m_{13}$	$-m_{12}$	$6 \cdot 4 \lambda_{1123}$	$6 \cdot 4 \lambda_{2123}$	$6 \cdot 4 \lambda_{3123}$
	$-m_{33}$	0	$-2 m_{13}$	$3 \cdot 4 \lambda_{1133}$	$3 \cdot 4 \lambda_{2133}$	$3 \cdot 4 \lambda_{3133}$
	0	$-3 m_{22}$	0	$4 \lambda_{1222}$	$4 \lambda_{2222}$	$4 \lambda_{3222}$
	0	$-2 m_{23}$	$-m_{22}$	$3 \cdot 4 \lambda_{1223}$	$3 \cdot 4 \lambda_{2223}$	$3 \cdot 4 \lambda_{3223}$
	0	$-m_{33}$	$-2 m_{23}$	$3 \cdot 4 \lambda_{1233}$	$3 \cdot 4 \lambda_{2233}$	$3 \cdot 4 \lambda_{3233}$
	0	0	$-3 m_{33}$	$4 \lambda_{1333}$	$4 \lambda_{2333}$	$4 \lambda_{3333}$
	$-4 m_{111}$	0	0	$5 \lambda_{11111}$	$5 \lambda_{21111}$	$5 \lambda_{31111}$
	$-3 m_{112}$	$-m_{111}$	0	$4 \cdot 5 \lambda_{11112}$	$4 \cdot 5 \lambda_{21112}$	$4 \cdot 5 \lambda_{31112}$
$m = 4$	$-3 m_{113}$	0	$-m_{111}$	$4 \cdot 5 \lambda_{11113}$	$4 \cdot 5 \lambda_{21113}$	$4 \cdot 5 \lambda_{31113}$
	$-2 m_{122}$	$-2 m_{112}$	0	$6 \cdot 5 \lambda_{11122}$	$6 \cdot 5 \lambda_{21122}$	$6 \cdot 5 \lambda_{31122}$
	$-2 m_{123}$	$-m_{113}$	$-m_{112}$	$12 \cdot 5 \lambda_{11123}$	$12 \cdot 5 \lambda_{21123}$	$12 \cdot 5 \lambda_{31123}$
	$-2 m_{133}$	0	$-2 m_{113}$	$6 \cdot 5 \lambda_{11133}$	$6 \cdot 5 \lambda_{21133}$	$6 \cdot 5 \lambda_{31133}$
	$-m_{222}$	$-3 m_{122}$	0	$4 \cdot 5 \lambda_{11222}$	$4 \cdot 5 \lambda_{21222}$	$4 \cdot 5 \lambda_{31222}$
	$-m_{223}$	$-2 m_{123}$	$-m_{122}$	$12 \cdot 5 \lambda_{11223}$	$12 \cdot 5 \lambda_{21223}$	$12 \cdot 5 \lambda_{31223}$
	$-m_{233}$	$-m_{133}$	$-2 m_{123}$	$12 \cdot 5 \lambda_{11233}$	$12 \cdot 5 \lambda_{21233}$	$12 \cdot 5 \lambda_{31233}$
	$-m_{333}$	0	$-3 m_{133}$	$4 \cdot 5 \lambda_{11333}$	$4 \cdot 5 \lambda_{21333}$	$4 \cdot 5 \lambda_{31333}$
	0	$-4 m_{222}$	0	$5 \lambda_{12222}$	$5 \lambda_{22222}$	$5 \lambda_{32222}$
	0	$-3 m_{223}$	$-m_{222}$	$4 \cdot 5 \lambda_{12223}$	$4 \cdot 5 \lambda_{22223}$	$4 \cdot 5 \lambda_{32223}$
	0	$-2 m_{233}$	$-2 m_{223}$	$6 \cdot 5 \lambda_{12233}$	$6 \cdot 5 \lambda_{22233}$	$6 \cdot 5 \lambda_{32233}$
	0	$-m_{333}$	$-3 m_{233}$	$4 \cdot 5 \lambda_{12333}$	$4 \cdot 5 \lambda_{22333}$	$4 \cdot 5 \lambda_{32333}$
	0	0	$-4 m_{333}$	$5 \lambda_{13333}$	$5 \lambda_{23333}$	$5 \lambda_{33333}$
	$-5 m_{1111}$	0	0	$6 \lambda_{111111}$	$6 \lambda_{211111}$	$6 \lambda_{311111}$
	$-4 m_{1112}$	$-m_{1111}$	0	$5 \cdot 6 \lambda_{111112}$	$5 \cdot 6 \lambda_{211112}$	$5 \cdot 6 \lambda_{311112}$
	$-4 m_{1113}$	0	$-m_{1111}$	$5 \cdot 6 \lambda_{111113}$	$5 \cdot 6 \lambda_{211113}$	$5 \cdot 6 \lambda_{311113}$
$m = 5$	$-3 m_{1122}$	$-2 m_{1112}$	0	$10 \cdot 6 \lambda_{111122}$	$10 \cdot 6 \lambda_{211122}$	$10 \cdot 6 \lambda_{311122}$
	$-3 m_{1123}$	$-m_{1113}$	$-m_{1112}$	$20 \cdot 6 \lambda_{111123}$	$20 \cdot 6 \lambda_{211123}$	$20 \cdot 6 \lambda_{311123}$
	$-3 m_{1133}$	0	$-2 m_{1113}$	$10 \cdot 6 \lambda_{111133}$	$10 \cdot 6 \lambda_{211133}$	$10 \cdot 6 \lambda_{311133}$
	$-2 m_{1222}$	$-3 m_{1122}$	0	$10 \cdot 6 \lambda_{111222}$	$10 \cdot 6 \lambda_{211222}$	$10 \cdot 6 \lambda_{311222}$
	$-2 m_{1223}$	$-2 m_{1123}$	$-m_{1122}$	$30 \cdot 6 \lambda_{111223}$	$30 \cdot 6 \lambda_{211223}$	$30 \cdot 6 \lambda_{311223}$
	$-2 m_{1233}$	$-m_{1133}$	$-2 m_{1123}$	$30 \cdot 6 \lambda_{111233}$	$30 \cdot 6 \lambda_{211233}$	$30 \cdot 6 \lambda_{311233}$
	$-2 m_{1333}$	0	$-3 m_{1133}$	$10 \cdot 6 \lambda_{111333}$	$10 \cdot 6 \lambda_{211333}$	$10 \cdot 6 \lambda_{311333}$
	$-m_{2222}$	$-4 m_{1222}$	0	$5 \cdot 6 \lambda_{112222}$	$5 \cdot 6 \lambda_{212222}$	$5 \cdot 6 \lambda_{312222}$
	$-m_{2223}$	$-3 m_{1223}$	$-m_{1222}$	$20 \cdot 6 \lambda_{112223}$	$20 \cdot 6 \lambda_{212223}$	$20 \cdot 6 \lambda_{312223}$
	$-m_{2233}$	$-2 m_{1233}$	$-2 m_{1223}$	$30 \cdot 6 \lambda_{112233}$	$30 \cdot 6 \lambda_{212233}$	$30 \cdot 6 \lambda_{312233}$
	$-m_{2333}$	$-m_{1333}$	$-3 m_{1233}$	$20 \cdot 6 \lambda_{112333}$	$20 \cdot 6 \lambda_{212333}$	$20 \cdot 6 \lambda_{312333}$
	$-m_{3333}$	0	$-4 m_{1333}$	$5 \cdot 6 \lambda_{113333}$	$5 \cdot 6 \lambda_{213333}$	$5 \cdot 6 \lambda_{313333}$
	0	$-5 m_{2222}$	0	$6 \lambda_{122222}$	$6 \lambda_{222222}$	$6 \lambda_{322222}$
	0	$-4 m_{2223}$	$-m_{2222}$	$5 \cdot 6 \lambda_{122223}$	$5 \cdot 6 \lambda_{222223}$	$5 \cdot 6 \lambda_{322223}$
	0	$-3 m_{2233}$	$-2 m_{2223}$	$10 \cdot 6 \lambda_{122233}$	$10 \cdot 6 \lambda_{222233}$	$10 \cdot 6 \lambda_{322233}$
	0	$-2 m_{2333}$	$-3 m_{2233}$	$10 \cdot 6 \lambda_{122333}$	$10 \cdot 6 \lambda_{222333}$	$10 \cdot 6 \lambda_{322333}$
	0	$-m_{3333}$	$-4 m_{2333}$	$5 \cdot 6 \lambda_{123333}$	$5 \cdot 6 \lambda_{223333}$	$5 \cdot 6 \lambda_{323333}$
	0	0	$-5 m_{3333}$	$6 \lambda_{133333}$	$6 \lambda_{233333}$	$6 \lambda_{333333}$
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

(32)

Table 2: The system of equations Eq. 30 can be grouped according to a three column matrix, in regard to their right hand side. A similar procedure is applied to the unknown coefficients of $bf\lambda_k$. Thus, the system can be expressed as $[\mathbf{A}, \mathbf{B}, \mathbf{C}] = \mathbf{G}_2 [\mathbf{a}, \mathbf{b}, \mathbf{c}]$.

Thus, by taking into account the symmetry of the coefficients of \mathcal{P}_n , we define, according to Table 2, the following three column matrices $\mathbf{X} = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$ and $\mathbf{Y} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$. Then, the following equality is satisfied

$$\mathbf{Y} = \mathbf{G}_2 \mathbf{X} \quad (33)$$

which is equivalent to the system of equations Eq. 30.

The coefficients to compute are the elements of the symmetric tensors λ_k , for orders $k = 1, \dots, n$ (since order zero corresponds to the normalisation factor). In total there are $\sum_{k=1}^n \binom{k+2}{2} = \binom{n+3}{3} - 1$ independent coefficients. For each column of matrix \mathbf{X} there are: one element of the symmetric tensor λ_1 , associated with the equations for $m = 0, 1, \dots, n-1$, which multiplies the moments of orders $0, 1, \dots, n-1$, respectively; three elements of λ_2 , which multiply the moments of orders $1, 2, \dots, n$; and, in general, $\binom{k+2}{2}$ elements of the symmetric tensor λ_{k+1} , which multiply the moments of orders $k, \dots, k+n-1$, up to reach the value $k = n-1$. The moments up to order $2(n-1)$ are involved. Thus, for example, for $n = 2$ we use the matrix \mathbf{G}_2 with the first row containing moments up to first order ($1+3=4$ columns in total), and the last row containing moments up to order 2, being a 4×4 matrix. For $n = 4$ we use the matrix \mathbf{G}_2 with the first row containing moments up to order 3 ($1+3+6+10=20$ columns in total), and the last row containing moments up to order 6, being a 20×20 matrix. Similarly, for $n = 6$ we use the matrix \mathbf{G}_2 with the first row containing moments up to order 5 ($1+3+6+10+15+21=56$ columns in total) and the last row containing moments up to order 10, being a 56×56 matrix.

5 Application

The system of equations Eq. 30 allows us to compute the coefficients of tensors λ_k , $1 \leq k \leq n$, in terms of the velocity moments up to order $2(n-1)$, which is the highest order involved in Eq. 28, as it is discussed in the previous section.

For the case $n = 2$, it is easy to solve the Gramian system in an analytical way, and to find how moments of order higher than two depend on the second ones (Appendix A). For higher values of n , however, it must be done by using the numerical procedure outlined in Appendix B. Also, for $n = 2$, the integrability of the distribution function in an infinite velocity domain is easily derived from the tensor λ_2 , since $\lambda_{ij} = -\frac{1}{2}\mu_{ij}^{-1}$, where the tensor of second central moments is positive definite. For $n \geq 4$, however, it is not possible to guarantee, in a general way, the definiteness of the tensor λ_n . This is a problem related to Hilbert's 17th problem, which is obviously out of the scope of the present work. However, by using a finite velocity domain, as wide as needed according to the working stellar sample, such a problem may be avoided under the conditions given for truncated distributions.

I discuss two illustrative cases. The first one is to test the ability of the maximum entropy procedure in reconstructing a truncated distribution associated with a velocity bounded stellar sample, Sample I, which is composed of a complex mixture of early-type and young disk stars. In the second example, a nearly complete and kinematically representative local sample, Sample II is used, which contains thin and thick disk populations. Both samples have been drawn from a reference sample used by Cubarsi & Alcobé (2004), which was obtained by crossing the Hipparcos Catalogue (ESA 1997) with radial velocities from the Hipparcos Input Catalogue (ESA 1992). In order to get a representative sample of the solar neighbourhood, the reference sample (13,678 stars) was limited to a trigonometric distance of 300 pc, where the only input data were the velocity components (U, V, W) in a cartesian heliocentric coordinates system, with U toward the Galactic

centre, V in the rotational direction, and W perpendicular to Galactic plane, positive in the direction of the North Galactic pole. Both subsamples had been studied by using other statistical methods, and they had well known kinematic features, which were consistent with similar studies of the solar neighbourhood. Hence it will be possible to compare the results of the present method with the previous ones.

- Sample I. It was the result of filtering the whole sample from a bimodal pattern named MEMPHIS algorithm (Alcobé & Cubarsi 2005) in order to detect discontinuities in the velocity distribution, which are associated with some differentiated behaviour of the stellar populations. Each discontinuity was detected from a drop of the entropy of the mixture probability as scanning the whole sample with the so-called sampling parameter. In this case, the sampling parameter led to an absolute value of the total space motion $|\mathbf{V}| \leq 51 \text{ Km s}^{-1}$. The sample was composed of 10,195 stars belonging to the core of thin disk, and, according to MEMPHIS algorithm, it was the more representative sample containing early-type and young disk stars, by having left aside most of the old disk stars.

For this bounded sample, neither the maximum entropy approach with $n = 2$, with moments up to second order, nor the approach with $n = 4$, with moments up to sixth order, are able to provide a realistic map of the truncated distribution, mainly due to the existing asymmetry along the radial velocity component. However, for $n = 6$, by fitting up to tenth moments, the maximum entropy approach gives a much improved portrait of actual asymmetries. The contourplots of the velocity distribution in each velocity plane are displayed in Fig. 1, as well as, in the last row, for $n = 6$, the sections of the distribution function (not normalised) for each velocity component. For such a worthily bounded sample, a sharp bell shaped distribution is obtained, although with very short wings, in particular for the U velocity component. Also notice the smoothing trend of the entropy approach.

It is therefore possible to estimate the truncated velocity distribution of younger stars from a maximum entropy function by using a six degree polynomial. The results are totally consistent with the contourplots obtained by Dehnen (1998) when inferring the velocity distribution of his total sample (AL), in particular for the innermost dark contour. Also, the shape of the velocity distribution that Skuljan et al. (1999) obtained for early-type stars is absolutely similar to ours. In particular, on the direction of the vertex deviation $V = 0.21U$, which may be obtained from the approach with $n = 2$, the distribution function has a clear bimodal pattern, as shown in Fig. 2. The lowest peak is placed around the Hyades stream, and the greater peak around the Sirius-UMa stream (Skuljan et al. 1999). Both estimations were also done by using Hipparcos data, although they used much more sophisticated methods to infer the velocity distribution.

- Sample II. This is the nearly complete sample, but in order to exclude a few stars with erroneous data and, perhaps a few stars belonging to the halo, which could be considered as sample noise, MEMPHIS algorithm provided an optimal value of the selecting parameter for $|\mathbf{V}| \leq 210 \text{ Km s}^{-1}$. The sample contains 13,530 stars and, as approaching the velocity boundary, the distribution vanishes. Therefore, to practical purposes the sample can be considered unbounded.

The maximum entropy procedure with $n = 2$ tries to represent the whole distribution from an unique ellipsoidal distribution. Of course, odd-order moments and even-order moments higher than four are not fitted. The approach with $n = 4$, by using up to sixth-order

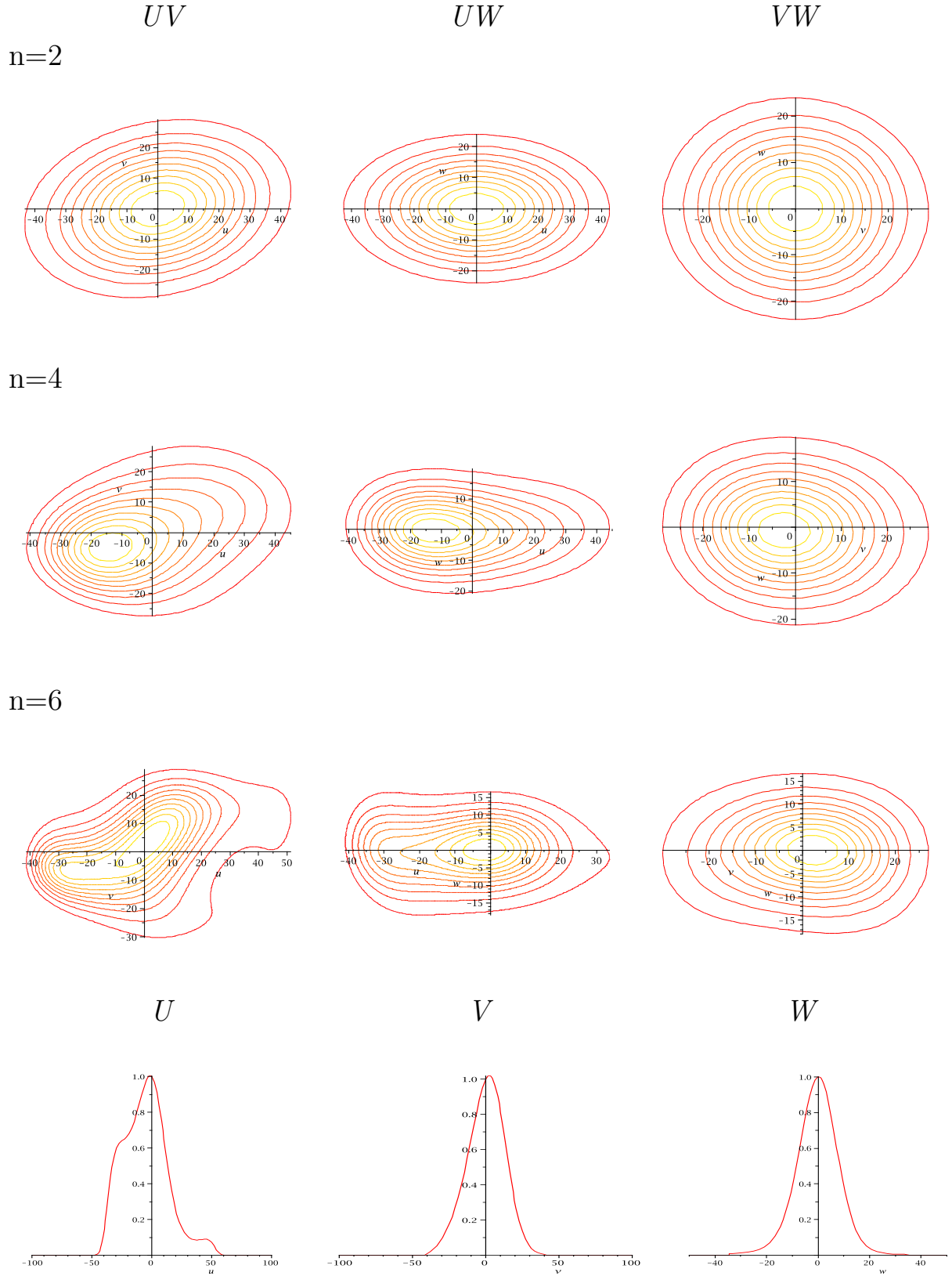


Figure 1: Contourplots of the velocity distribution for stars with $|V| \leq 51 \text{ km s}^{-1}$, Sample I, in terms of the peculiar velocities. The approach $n=2$ uses moments up to second order, $n=4$ up to sixth order, and $n=6$ up to tenth order. The strong asymmetry of the velocity distribution, mainly in the U velocity component, may be sufficiently described for the case $n=6$, by fitting moments up to tenth order. In the last row, for $n=6$, the sections of the velocity distribution function (non normalised) for each velocity component are displayed.

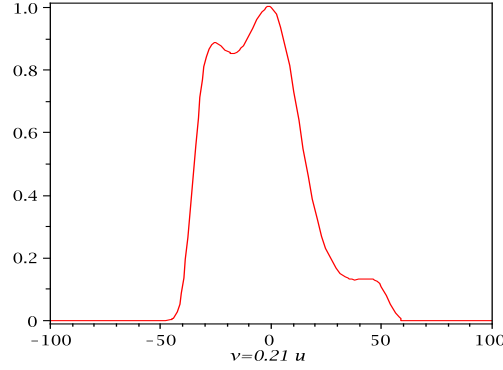


Figure 2: For $n = 6$, the maximum entropy distribution (non normalised) shows a net bimodal shape in the direction of the vertex deviation $V = 0.21U$, obtained from the approach with $n = 2$. The lowest peak is placed around the Hyades stream, and the greater peak around the Sirius-UMa stream.

moments, gives a clearly improved result. Finally, the approach with $n = 6$ is also computed, in order to be compare with the previous one, by fitting up to tenth moments. Differences between $n = 4$ and $n = 6$ are not significant at all. For each approach, the contourplots of the velocity distribution in each velocity plane are displayed in Fig. 3, as well as the sections of the distribution function for each velocity component.

The main features of the resulting maximum entropy distribution are also similar to those described by Dehnen (1998) for the background distribution of a kinematically representative local sample. They can be easily deduced from Fig. 3, either for $n = 4$ or $n = 6$.

- The core distribution has a radial velocity similar to the mean of the whole sample, but a rotational velocity greater than the mean.
- All the distribution has a similar and clear vertex deviation on the plane UV , and no deviation on other planes.
- $W = 0$ is basically a symmetry plane.
- There is some skewness on the plane VW .
- As approaching the wings of the distribution, the contourplots become more centred around the rotation mean velocity, which indicates a shift towards lower rotational velocities of the background with respect to the core distribution.

6 Discussion

The generic trends of the complete local velocity distribution shown by the entropy approach are also consistent with those of more recent works (Veltz et al. 2008, Vallerani et al. 2006, Famaey et al. 2005, Skuljan et al. 1999) where they can be compared. For example, the constant vertex deviation of the contourplots as approaching the wings of the distribution, and the axis ratio of the ellipsoidal background. The resulting velocity distribution, according to the most significant polynomial coefficients, can be expressed as a product of two exponential functions in the form

$$f = \varphi_1(Q) \varphi_2(h) \quad (34)$$

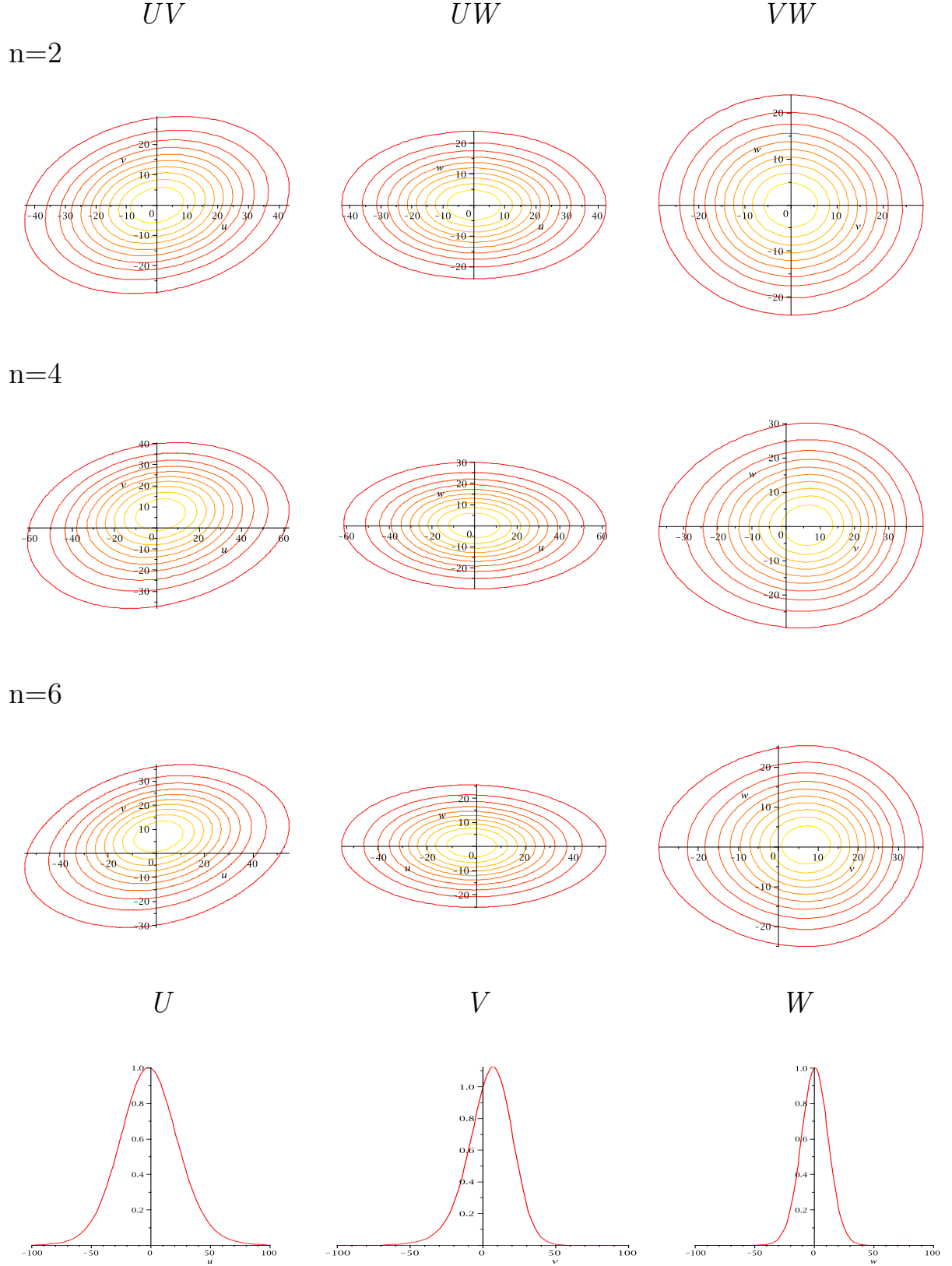


Figure 3: Contourplots of the local velocity distribution in terms of the peculiar velocities for Sample II, which is a kinematically representative sample of the solar neighbourhood. Differences between $n=4$ and $n=6$ are not significant, while the approach $n=2$ does explain the slight asymmetry of the distribution. In the last row, for the case $n=6$, the sections of the velocity distribution function (non normalised) for each velocity component are displayed.

where Q is a quadratic positive definite form, which gives the background ellipsoidal shape of the distribution, with axis ratio about 1:0.6:0.5, a symmetry plane $W = 0$, as expected for disk stellar samples, and an overall vertex deviation in the UV velocity components of about 12° .

The function $\varphi_2(h)$, which can be expressed in terms of the angular momentum integral h , is a perturbation factor, which is quadratic in the V velocity alone, and gives account of the skewness and of the slight shift of the velocity ellipsoids in terms of the rotation velocity.

However, since the maximum entropy approach takes the distribution as a whole, and tends to smooth the existing discontinuities, it is not possible a detailed comparison with mixture models but for general trends. For example, the same total Sample II was used by Alcobé & Cubarsi (2005) and Cubarsi & Alcobé (2006), to describe some differentiated kinematic behaviours of the neighbourhood stars. The whole sample was successfully described from a superposition of two trivariate Schwarzschild distributions, associated with thin and thick disk populations. The shape of the resulting maximum entropy distribution is, as expected, quite similar to the one of the prominent thin disk component. However, within the thin disk component, the gaussian mixture approach was unable to describe the so-called early-type and young disk stars mixture. On the contrary, the entropy approach do give a suitable estimation for those stars in Sample I, so that it can be used as an alternative approach to model multimodal distributions. Of course those stars are not in thermodynamical equilibrium, however, from a numerical viewpoint, it could be said that, lacking a more detailed knowledge of the underlying microscopical details, a maximum entropy distribution could be very close to the real distribution, although rather different analytically.

In regard to the method itself, the following aspects may be summarised. The maximum entropy approach, although successfully applied to a wide disparity of actual problems, has been rarely used to solve the classical moment problem of stellar kinematics. Instead, a number of statistical techniques, maximum likelihood-based multivariate sampling algorithms, wavelets-based algorithms, among others, had been proved more appropriate than the moments method to give an accurate map of the local stellar velocity distribution. The moment approach had basically two difficulties in front of other methods: The low accuracy of available data, and the complexity of the trivariate model, as in the case of mixture distributions, in order to estimate the parameters. However, from large stellar catalogues it is nowadays possible to compute with a sufficient degree of accuracy a reasonable number of moments. Then, although with the minimum set of moment constraints the parameter estimation for a maximum entropy function requires some complex computational procedures, if an extended set of moments is available, the current method provides a linear estimation algorithm, which is given by a Gramian system of equations, that leads to a fast and suitable estimation of the velocity distribution, as shown in both case examples.

Appendix A

The Gramian system and the moment recurrence can be straightforward solved for the case $n = 2$, which corresponds to a Schwarzschild distribution. For the sake of simplicity, and without losing generality, we use the central moments μ_n , so that $\mu_1 = \mathbf{0}$. Then, the equations Eq. 26 and Eq. 27, for $m = 0, 1$, become

$$\lambda_k = 0, \quad k = 1, 2, 3 \quad (35)$$

and

$$\delta_{ik} + 2\lambda_{jk}\mu_{ji} = 0 \implies \lambda_{ik} = -\frac{1}{2}\mu_{ik}^{-1}, \quad i, k = 1, 2, 3 \quad (36)$$

where μ_{ik}^{-1} are the elements of the inverse of the covariance matrix.

Therefore, above relation shows that the tensor of elements λ_{ik} is a definite negative form, and it leads to an integrable distribution function.

Now we can apply the same procedure for $m \geq 2$, in order to obtain higher-order moments in terms of the second moments. Thus, for $m = 2$, according to Eq. 28, with $m = n - 1$, and bearing in mind Eq. 35, we have

$$\lambda_k\mu_{ij} + 2\lambda_{lk}\mu_{jil} = 0 \implies \mu_{ijk} = 0, \quad i, j, k = 1, 2, 3 \quad (37)$$

The result accounts for the obvious symmetry of the distribution, with vanishing odd-order central moments.

Similarly, for $m = 3$, we get

$$\delta_{i_1 i_4} \mu_{i_2 i_3} + \delta_{i_2 i_4} \mu_{i_1 i_3} + \delta_{i_3 i_4} \mu_{i_1 i_2} + \lambda_{i_4} \mu_{i_1 i_2 i_3} + 2\lambda_{j i_4} \mu_{j i_1 i_2 i_3} = 0, \quad i_1, i_2, i_3, i_4 = 1, 2, 3 \quad (38)$$

Then, taking into account Eq. 36, we multiply by $\mu_{k i_4}$. Since the third moments are null, by reordering indices we obtain the following moment recurrence relation

$$\mu_{i_1 i_2 i_3 i_4} = \mu_{i_1 i_4} \mu_{i_2 i_3} + \mu_{i_2 i_4} \mu_{i_1 i_3} + \mu_{i_3 i_4} \mu_{i_1 i_2} \quad (39)$$

Above relationship is the well known property a gaussian distribution, which characterise it from having vanishing fourth cumulants.

And, in general, according to Eq. 28, for m even, we obtain a vanishing set of odd-order central moments, and, for m odd, we obtain the relation

$$\delta_{i_1 i_{m+1}} \mu_{i_2 \dots i_m} + \dots + \delta_{i_j i_{m+1}} \mu_{i_1 \dots \widehat{i_j} \dots i_m} + \dots + \delta_{i_m i_{m+1}} \mu_{i_1 \dots i_{m-1}} + 2\lambda_{j i_{m+1}} \mu_{j i_1 \dots i_m} = 0 \quad (40)$$

Once again, by multiplying by $\mu_{k i_{m+1}}$, and by reordering indices, we get

$$\mu_{i_1 \dots i_m i_{m+1}} = \mu_{i_1 i_{m+1}} \mu_{i_2 \dots i_m} + \dots + \mu_{i_j i_{m+1}} \mu_{i_1 \dots \widehat{i_j} \dots i_m} + \dots + \mu_{i_m i_{m+1}} \mu_{i_1 \dots i_{m-1}} \quad (41)$$

This is the general moment recurrence relationship for trivariate normal distributions, which leads to a vanishing set higher-order cumulants.

Appendix B

In this appendix the details to invert the Gramian system of equations are given. First we write the system matrix \mathbf{G}_2 of Table 1 in greek indices notation, namely $\mathbf{\Gamma}$. In such a notation it can

be easily interpreted as a matrix of inner products, which is associated with a finite set of linearly independent vectors

$$\phi_{\alpha\beta\gamma} = V_1^\alpha V_2^\beta V_3^\gamma \quad (42)$$

belonging to a basis of the space of square-integrable functions in Γ_V , with respect to the weight function $e^{\mathcal{P}_n}$, so that

$$\Gamma(\alpha, \beta, \gamma; \lambda, \mu, \nu) = \langle \phi_{\alpha\beta\gamma}, \phi_{\lambda\mu\nu} \rangle \quad (43)$$

Starting from Eq. 13, Eq. 22 can be written by using greek indices notation.¹ By taking the V_1 derivative we have

$$\begin{aligned} \frac{\partial(V_1^\alpha V_2^\beta V_3^\gamma e^{\mathcal{P}_n})}{\partial V_1} &= \alpha V_1^{\alpha-1} V_2^\beta V_3^\gamma e^{\mathcal{P}_n} + V_1^\alpha V_2^\beta V_3^\gamma \frac{\partial \mathcal{P}_n}{\partial V_1} e^{\mathcal{P}_n} = \\ &= \left(\alpha V_1^{\alpha-1} V_2^\beta V_3^\gamma + V_1^\alpha V_2^\beta V_3^\gamma \sum_{k=0}^n \sum_{\iota+\mu+\nu=k} \lambda_{\iota\mu\nu}^{(k)} V_1^{\iota-1} V_2^\mu V_3^\nu \right) e^{\mathcal{P}_n} \end{aligned}$$

Since the last summation can be carried out from $\iota \geq 1$ instead of $\iota \geq 0$, by renaming $\iota - 1$ as ι we can then write

$$\frac{\partial(V_1^\alpha V_2^\beta V_3^\gamma e^{\mathcal{P}_n})}{\partial V_1} = \left(\alpha V_1^{\alpha-1} V_2^\beta V_3^\gamma + V_1^\alpha V_2^\beta V_3^\gamma \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \lambda_{(\iota+1)\mu\nu}^{(k+1)} (\iota+1) V_1^\iota V_2^\mu V_3^\nu \right) e^{\mathcal{P}_n} \quad (44)$$

Similarly, for the other derivatives, we have

$$\begin{aligned} \frac{\partial(V_1^\alpha V_2^\beta V_3^\gamma e^{\mathcal{P}_n})}{\partial V_2} &= \left(\beta V_1^\alpha V_2^{\beta-1} V_3^\gamma + V_1^\alpha V_2^\beta V_3^\gamma \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \lambda_{\iota(\mu+1)\nu}^{(k+1)} (\mu+1) V_1^\iota V_2^\mu V_3^\nu \right) e^{\mathcal{P}_n} \\ \frac{\partial(V_1^\alpha V_2^\beta V_3^\gamma e^{\mathcal{P}_n})}{\partial V_3} &= \left(\gamma V_1^\alpha V_2^\beta V_3^{\gamma-1} + V_1^\alpha V_2^\beta V_3^\gamma \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \lambda_{\iota\mu(\nu+1)}^{(k+1)} (\nu+1) V_1^\iota V_2^\mu V_3^\nu \right) e^{\mathcal{P}_n} \end{aligned} \quad (45)$$

Above expressions can be substituted into Eq. 23 and, by using the notation introduced in Eq. 43, we get the following set of equations.

¹From a numerical viewpoint, and in order to establish an exact correspondence between greek and latin indices, the summation terms of Eq. 13 are ordered as follows. Any quantity $\varphi_{\alpha\beta\gamma}^{(k)}$ written in greek notation, with $\alpha + \beta + \gamma = k$, and each index taking values from 0 to k , is sorted by associating the set of indices with a number $[\alpha\beta\gamma]_k$, which is the expression of “ $\alpha\beta\gamma$ ”, according to the numeral system in base $(k+1)$. Then, to simplify, we can write $\varphi_{\alpha\beta\gamma}^{(k)} = \varphi_{[\alpha\beta\gamma]_k}$ in terms of one index alone, by taking the summation terms as decreasing values of $[\alpha\beta\gamma]_k$, with $0 \leq \gamma \leq 4$, $4 - \gamma \leq \beta \leq 4$ and $4 - \gamma - \beta \leq \alpha \leq 4$. Thus, in Eq. 13, \mathcal{P}_n can be written, bearing in mind Eq. 42, as depending on one summation index,

$$P_n = \sum_{[\alpha\beta\gamma]_k} \lambda_{[\alpha\beta\gamma]_k} \phi_{[\alpha\beta\gamma]_k}$$

For example, for $k = 4$ the ordination of greek indices is 400, 310, 301, 220, 211, 202, 130, 121, 112, 103, 040, 031, 022, 013, 004. Above expression and Eq. 14, in latin indices, have the same summation order. Hereafter, such an indexation order will be assumed when working with greek indices, although, for the sake of simplicity, the former notation is maintained.

$$\begin{aligned}
A_{\alpha\beta\gamma}^{(m)} &\equiv -\alpha m_{(\alpha-1)\beta\gamma} = \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma(\alpha, \beta, \gamma; \iota, \mu, \nu) (\iota+1) \lambda_{(\iota+1)\mu\nu}^{(k+1)} \\
B_{\alpha\beta\gamma}^{(m)} &\equiv -\beta m_{\alpha(\beta-1)\gamma} = \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma(\alpha, \beta, \gamma; \iota, \mu, \nu) (\mu+1) \lambda_{\iota(\mu+1)\nu}^{(k+1)} \\
C_{\alpha\beta\gamma}^{(m)} &\equiv -\gamma m_{\alpha\beta(\gamma-1)} = \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma(\alpha, \beta, \gamma; \iota, \mu, \nu) (\nu+1) \lambda_{\iota\mu(\nu+1)}^{(k+1)}
\end{aligned} \tag{46}$$

which stands for $\alpha + \beta + \gamma = m$ and $0 \leq m \leq n-1$. Notice that, according to this notation, all the moments having a negative index must be considered null.

The left hand side terms of the foregoing expressions are related to the three-column matrix \mathbf{Y} of §4, although, they were there written in latin notation, as

$$[A_i, B_i, C_i] = [A_{\alpha\beta\gamma}^{(m)}, B_{\alpha\beta\gamma}^{(m)}, C_{\alpha\beta\gamma}^{(m)}] \tag{47}$$

Similarly, the three-column matrix \mathbf{X} , also in latin notation in §4, has now its corresponding elements, according to

$$[a_i, b_i, c_i] = [(\iota+1) \lambda_{(\iota+1)\mu\nu}^{(k+1)}, (\mu+1) \lambda_{\iota(\mu+1)\nu}^{(k+1)}, (\nu+1) \lambda_{\iota\mu(\nu+1)}^{(k+1)}] \tag{48}$$

Therefore, the coefficients of tensors λ_k , $1 \leq k \leq n$, involved in Eq. 13, other than the normalisation factor λ_0 , can be explicitly obtained in terms of the moments up to order $2(n-1)$ as

$$\begin{aligned}
\lambda_{(\alpha+1)\beta\gamma}^{(l+1)} &= \frac{-1}{\alpha+1} \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma^{-1}(\alpha, \beta, \gamma; \iota, \mu, \nu) \iota m_{(\iota-1)\mu\nu} \\
\lambda_{\alpha(\beta+1)\gamma}^{(l+1)} &= \frac{-1}{\beta+1} \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma^{-1}(\alpha, \beta, \gamma; \iota, \mu, \nu) \mu m_{\iota(\mu-1)\nu} \\
\lambda_{\alpha\beta(\gamma+1)}^{(l+1)} &= \frac{-1}{\gamma+1} \sum_{k=0}^{n-1} \sum_{\iota+\mu+\nu=k} \Gamma^{-1}(\alpha, \beta, \gamma; \iota, \mu, \nu) \nu m_{\iota\mu(\nu-1)}
\end{aligned} \tag{49}$$

for $l = 0, \dots, n-1$, through $\Gamma^{-1}(\alpha, \beta, \gamma; \iota, \mu, \nu)$, which is the corresponding element of the inverse of the matrix $\mathbf{\Gamma}$.

References

- Alcobé S., Cubarsi R. 2005, A&A 442, 929
- Cercignani, C. 1988, The Boltzmann Equation and its Applications. Springer-Verlag, Berlin
- Chandrasekhar S. 1942, Principles of Stellar Dynamics. Dover Publications Inc., New York
- Charlier, C.W.L. 1926, The Motion and the Distribution of the Stars. University of California, Berkeley
- Cubarsi R., Alcobé S. 2004, A&A 427, 131
- Cubarsi R., Alcobé S. 2006, A&A 457, 537
- Cubarsi R. 2007, MNRAS 207, 380
- de Bruin R., Salomé D., Schaafsma W. 1999, Comput. Statist. Data Anal. 30, 19
- Dehnen W. 1998, AJ 115, 2384
- Dejonghe H. 1983, A&A 133, 225
- Erickson R.R. 1975, ApJ 195, 343
- ESA 1992, The Hipparcos Input Catalog. ESA SP-1136
- ESA 1997, The Hipparcos Catalogue. ESA SP-1200
- Evans N.W., Haefner R.M., de Zeeuw P.T. 1997, MNRAS 286, 315
- Famaey B., Van Caelenberg K., Dejonghe H. 2002, MNRAS 335, 201
- Famaey, B., Jorissen, A., Luri, X., Mayor, M., Udry, S., Dejonghe, H., Turon, C. 2005, A&A, 430, 165
- Fricke W. 1952, Astron. Nachr. 280, 193
- Gerhard O.E. 1993, MNRAS 265, 213
- Hénon M. 1973, A&A 24, 229
- Jaynes E.T. 1957, Phys. Rev. 106, 620 & 108, 171
- Jaworsky W. 1987, J Phys. A: Math. Gen. 20, 915
- Jiang Z., Ossipkov L. 2007, MNRAS 379, 1133
- Kagan A.M., Linnik Y.V., Rao C.R. 1973, Characterization Problems in Mathematical Statistics. Wiley, New York
- Kapteyn J.C. 1922, ApJ 55, 302
- Kormendy J., Bender R., Evans A.S., Richstone D. 1998, AJ 115, 1823
- Kouskoulas Y., Pierce L.E., Ulaby F.T. 2004, IEEE Trans. Geosci. Rem. Sens. 42, 457
- Lynden-Bell D. 1967, MNRAS 136, 101
- Mead L.R., Papanicolaou N. 1984, J. Math. Phys. 25, 2405
- Skuljan, J., Hearnshaw, J.B., Cottrell, P.L. 1999, MNRAS, 308, 731
- Soubiran C., Girard, P. 2005, A&A 438, 139
- Strömberg, G. 1925, ApJ 61, 363
- Stuart A., Ord J.K. 1987, Kendall's Advanced Theory of Statistics, Vol. 1, Distribution Theory. Ch. Griffin & Co., London
- Vallenari A., Pasetto S., Bertelli G., Chiosi C., Spagna A., Lattanzi M. 2006, AA 451, 125
- van der Marel R.P., Franx M. 1993, ApJ 407, 525
- Veltz et al. 2008, A&A 480, 753
- White S.D.M. 1985, ApJ 294, L99